Reviewer’s report

Title: The association between neighborhood greenness and cardiovascular disease: an observational study

Version: 2 Date: 26 March 2012

Reviewer: Janice Bell

Reviewer’s report:

This is a well-written paper that contributes to the emerging literature focused on relations between neighborhood greenness and health. The primary contributions are: a) the novel focus on cardiovascular disease as an outcome; and b) the specification of NDVI including its mean as well as its variability.

The paper would be strengthened by addressing the following:

Major Compulsory Revisions

1. There is no consensus in the literature about the size of spatial unit in which greenness ought to be measured. What was the rationale for choosing a 1600 m service area? In what ways might the size of the spatial unit have influenced the study findings or their interpretation?

2. Provide summary information (similar to that given in Table 1) for the NDVI variable across the study setting, including the mean and the range. Can you also interpret these details to assist the reader to understand the overall greenness of the study setting?

3. Justify why the NDVI variable treated categorically rather than continuously. What was the rationale for the cut-points that were chosen?

4. Discuss the rationale for the choice of control variables. A rationale was given for air pollutants (lines 112 – 118) but not for the other variables. Total minutes of walking per week may be over-controlling given that physical activity was given as the hypothesized mechanism for a greenness-CVD relationship. Could physical activity be examined as an outcome?

5. To the paragraph in lines 112 – 118 outlining the rationale for including air pollutants, add the description of the variable(s) and data source(s) used in the current study to measure “environmental factors”.

6. What was done to measure and account for possible spatial autocorrelation in the statistical models? What were the results?

7. Lines 138 – 143. Please clarify here, and in Table 3, whether the association with mean greenness controlled for variability in greenness and vice versa. The footnote in the table referring to mean greenness under “environmental factors” is confusing.
8. Present additional results for the association of a) mean greenness (without controlling for variability) with the study outcomes; and for variability (without controlling for mean) with the study outcomes. (If these are among the models already presented in Tables 2 and 3, they should be clearly labeled accordingly.)

9. In the discussion, can you speculate further and provide any explanation(s) for the result presented in line 141 – 143?

10. The possibility of confounding by SES is important and could be elaborated. Also, provide citations or other justification for the claim that SES would be expected to be more highly correlated with levels of greenness than with variability. (There are a lot of non-green elements that might lead to high variability in high SES neighborhoods e.g., swimming pools, tennis courts, big homes).

11. Include the sample sizes in Tables 2 – 4. Also, clarify in text which analysis was performed on the full sample (n = 11,406) versus the subsample who granted permission for data linkage (74%). (Table 1 suggests the full sample was used for all analysis.) Did the subsample differ from those in the full sample on any of the key study variables?

Minor Essential Revisions

1. It would be helpful if the variables described in lines 107 – 111 were directly mapped in text to the labels "socio-demographic factors, biological factors, behavioral factors and environmental factors" given in lines 119 – 120 before these labels were introduced.

2. Can you also help the reader to understand how to interpret the findings for mean NDVI controlling for its variability? And for NDVI variability, controlling for mean NDVI? The figures are helpful for considering variability on its own—but it does not appear that variability was modeled as independent variable without mean greenness.

3. Are you able to sort what the mix or variability of greenness means in the data you have on hand—or at least speculate further for the reader about why “the mix of greenness is more relevant than the extent of greenness”? For instance, do the areas with higher variability include parks (greenness) that are accessible to cars (with parking lots)? Cycling or jogging paths along green settings? Bricks and mortar facilities for physical activity with outside landscaped recreation areas?

4. It is odd that the highest NDVI tertile is not significant across all models in Table 2 while the moderate category shows at least a weak association. Again, providing an explanation for the choice of NDVI cut-points would be helpful as well as presenting the findings with mean NDVI is modeled without NDVI variability and vice versa. The discussion should also include an explanation for the unexpected finding about why the highest category of NDVI might not be significant while the moderate category is significant.
Minor Issues not for Publication
1. Consider rewording the sentence in lines 39 – 40.
2. Consider rewording (perhaps dividing) the sentence in lines 64 – 68.
3. Sentence beginning line 148, “odds” are plural (i.e., should read odds were lower)

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.